A DOCPHOENIX

Raw Sequence Listing Error Summary

ERROR DETECTED   SUGGESTED CORRECTION

ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY PTO SOFTWARE

1 _____ Wrapped Nucleics

The number/text at the end of each line "wrapped" down to the next line.
This may occur if your file was retrieved in a word processor after creating it.
Please adjust your right margin to .3, as this will prevent "wrapping".

2 _____ Wrapped Aminos

The amino acid number/text at the end of each line "wrapped " down to the next line.
This may occur if your file was retrieved in a word processor after creating it.
Please adjust your right margin to .3, as this will prevent "wrapping".

3 _____ Incorrect Line Length

The rules require that a line not exceed 72 characters in length. This includes spaces.

4 _____ Misaligned Amino Acid Numbering

The numbering under each 5th amino acid is misaligned. This may be caused by the use of tabs between the numbering. It is recommended to delete any tabs and use spacing between the numbers.

5 _____ Non-ASCII

This file was not saved in ASCII (DOS) text, as required by the Sequence Rules.
Please ensure your subsequent submission is saved in ASCII text so that it can be processed.

6 _____ Variable Length

Sequence(s) _____ contain n's or Xaa's which represented more than one residue.
As per the rules, each n or Xaa can only represent a single residue.
Please present the maximum number of each residue having variable length and
indicate in the (ix) feature section that some may be missing.

7 _____ PatentIn ver. 2.0 "bug"

A "bug" in PatentIn version 2.0 has caused the <220>-<223> section to be missing from amino acid sequence(s) _____. Normally, PatentIn would automatically generate this section from the previously coded nucleic acid sequence. Please manually copy the relevant <220>-<223> section to the subsequent amino acid sequence. **This applies primarily to the mandatory <220>-<223> sections for Artificial or Unknown sequences.**

8 _____ Skipped Sequences (OLD RULES)

Sequence(s) _____ missing. If intentional, please use the following format for each skipped sequence:
(2) INFORMATION FOR SEQ ID NO:X:
(i) SEQUENCE CHARACTERISTICS:(Do not insert any headings under "SEQUENCE CHARACTERISTICS")
(xi) SEQUENCE DESCRIPTION:SEQ ID NO:X:
This sequence is intentionally skipped

Please also adjust the "(iii) NUMBER OF SEQUENCES:" response to include the skipped sequence(s).

9 _____ Skipped Sequences (NEW RULES)

Sequence(s) _____ missing. If intentional, please use the following format for each skipped sequence.
<210> sequence id number
<400> sequence id number
000

10 _____ Use of n's or Xaa's (NEW RULES)

Use of n's and/or Xaa's have been detected in the Sequence Listing.
Use of <220> to <223> is MANDATORY if n's or Xaa's are present.
In <220> to <223> section, please explain location of n or Xaa, and which residue n or Xaa represents.

11 _____ Use of <213>Organism (NEW RULES)

Sequence(s) _____ are missing this mandatory field or its response.

12 _____ Use of <220>Feature (NEW RULES)

Sequence(s) _____ are missing the <220>Feature and associated headings.
Use of <220> to <223> is MANDATORY if <213>ORGANISM is "Artificial" or "Unknown"
Please explain source of genetic material in <220> to <223> section.
(See "Federal Register," 6/01/98, Vol. 63, No. 104, pp. 29631-32)   (Sec. 1.823 of new Rules)

13 _____ PatentIn ver. 2.0 "bug"

Please do not use "Copy to Disk" function of PatentIn version 2.0. This causes a corrupted file, resulting in missing mandatory numeric identifiers and responses (as indicated on raw sequence listing).
Instead, please use "File Manager" or any other means to copy file to floppy disk.

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/701,626

DATE: 01/30/2001
TIME: 11:15:33

Input Set : A:\Neb-165a.app
Output Set: N:\CRF3\01302001\I701626.raw

**Does Not Comply
Corrected Diskette Needed**

PP.6-9

```
 3 <110> APPLICANT: New England Biolabs, Inc.
 4        Vaisvila, Romualdas
 5        Morgan, Richard D.
 6        Raleigh, Elisabeth
 8 <120> TITLE OF INVENTION: Restriction Enzyme Gene Discovery Method
10 <130> FILE REFERENCE: Gene Discovery Method
12 <140> CURRENT APPLICATION NUMBER: US/09/701,626
13 <141> CURRENT FILING DATE: 2000-01-12
15 <150> PRIOR APPLICATION NUMBER: 60/089,086
16 <151> PRIOR FILING DATE: 1998-06-12
18 <150> PRIOR APPLICATION NUMBER: 60/089,101
19 <151> PRIOR FILING DATE: 1998-06-12
21 <160> NUMBER OF SEQ ID NOS: 130
23 <170> SOFTWARE: PatentIn Ver. 2.0
25 <210> SEQ ID NO: 1
26 <211> LENGTH: 14143
27 <212> TYPE: DNA
28 <213> ORGANISM: Unknown
30 <220> FEATURE:
31 <223> OTHER INFORMATION: Genomic DNA of Pseudomonas Alcaligenes NEB#585
32        (ATCC 55044)
34 <400> SEQUENCE: 1
35 atcgatcagc cagacttttc gcacacgggc ggaccttggg cgagtcagcg ctatggttgg 60
36 ccgctgtggg ttgtcagtgc ccgtacgcgc aatctgtttc tttcgcaggg catgtccggc 120
37 tgggcgttcc ggcccgttct ggtcaccgac tcggctctct atgagcgcta tctcgctcta 180
38 agtcaggaac tttgcgcact gcttcgtgat gcaccgcaga gcaagctcga agaccgtgat 240
39 tggtaagcgg gggctattcg atcagtctcg gagcgaccaa actccagaaa cgacaaggcc 300
40 ctgaaaaaaa agcagggctt cgtctttgcg ggcgaatgga atcggacctc tttccgcctc 360
41 tgcatgtaac tggtctttgt ttgccaaatc tgcctatctc atgccggcca tgttggccag 420
42 tgcctgcatc atttggcctt tggtttcgac actttttcga cagccctgct agacatccct 480
43 ccctctgccc tcgtaacttc tgttccgatg gtgtcgcttg gcactatggt cttgtcgagt 540
44 gtcgcttttc atccagccta atgccgcgat tgcctcgctg agctgtagct gaatcaagga 600
45 cttagcggac gacaaggaat gttatgcgaa acatgtggcg gaataaatta cgccgcatgt 660
46 ttcgtctact tatagttagg ctacatatga gaatcagcgc agaccagctt gctcaagaat 720
47 cactgactga gttcggcgtg ctggcggcta agcttctggc aacgcgagag cttagccagt 780
48 tgtccgagaa gtttgggtat gcactggcct tcggaaggga accggcggct gccatagctg 840
49 aggaccttgc taggtgcttg tgcggacaaa atgcttcgcc ggcatctgaa taccccaaaa 900
50 tcaccgttaa gtatttcaag gaaaacgaaa gtagtctgtt ggcactcgta gagtgttatg 960
51 tacaaatgac cgcaagcgca aacattcttt tagagtctggt tgccgcacga aatggagagg 1020
52 caataaatct gtatctagaa ggcttgagtg ttgtagccta acaatgcgct caaagcgctc 1080
53 acttcgttcg ctgggaccgg cgaagccggc cccttagctt aatcgttaga aaccatcatg 1140
54 gataactggt acaacaccat cgaataccaa acccatgtag ccgaaaaact agaggcactt 1200
55 ggagaaacaa agtacgaccg cgaggcttat gaattcgcgc tagaggcata ccagtcgcg 1260
56 cctgaatatc atgaaaatat tcccacgccg cctctcaatc ttgggctcgc gtaccatgta 1320
57 agcgccttca acttcgcaca ctgctatgta cttcacgcta aagaagtgtt tgaagctcca 1380
58 aaagacacac tgagctcctg gggcgtattt tcctcaacgg acattggtga aattgtttat 1440
59 ggtttagtcc gtattggctt gctggaccaa ggccccgaag acaaaaaaga gcagtttgaa 1500
```

RAW SEQUENCE LISTING   DATE: 01/30/2001
PATENT APPLICATION: US/09/701,626   TIME: 11:15:33

Input Set : A:\Neb-165a.app
Output Set: N:\CRF3\01302001\I701626.raw

```
 60 gggttgtttt taatcaccga cgtgctgtga tgtcttctaa ctactggttc aagtcgttcg 1560
 61 cttcgctcac tcgggaccgg ctaaagccgg ccccttaacc aaacgttagc cacctcacga 1620
 62 agatttggag cccgcgtgaa caaagtcgat acaaacaaaa ttaaaacgga ttttctcggca 1680
 63 cgaattgatg aaaaaagagc gtggtttgat cgtatggcta cgcttataag cgggacaaac 1740
 64 accgagttaa ccgaccttaa ttttctttgc gagaactata taacatcaat atacgtagag 1800
 65 ctcgaatgct taatatcaga tttatttcat ggctacataa ataacaacaa caagacctac 1860
 66 atggcgcaca ttcaatcaaa aatcaagaac tccataactg acaagtactc tgcatggcac 1920
 67 gccacccata caacattcgc aggtccagag catattaatt cagcacagct cagcacgctc 1980
 68 cttgatccaa caagctggaa catcacattt aaagacgttt ccgcaatgaa agtacgagca 2040
 69 aaggaatacc tttcctcagt acacgaaaaa agattttcag gtatatctgc atccgatgga 2100
 70 gctcttattg atgccgcaca tgcaatcaga aattgcattg cacacaacag cgaaagctcc 2160
 71 aqaaagqtta tgaacaccaa aattaaaagc ttaattacag gcccagcttg ctcaaatgtc 2220
 72 ggccttgaac tcaccacaaa tagtgtgacc aaaataggaa agtatctccg tgcaaatgct 2280
 73 cagcaaagca tgcgagtgct gatttactca gatcgaataa aatctatcgg cctaagctta 2340
 74 taagtgtggg ctaacaatgc gctcaactgt cgctcacttc gttcgctgga cagccaaaag 2400
 75 ctacgctttt gtctgcccgt tagcttaatc gttaggaggc tctgcatgac tcgtgcaaca 2460
 76 gacaggttcg aagagcttct gcaatcacat gagttctcag ggcatattat tcgttgggtt 2520
 77 gcgatattcg aaggccgtct tgacggtgtg ttatcagttc attttctggg acttgaaagc 2580
 78 acctatgaat tctacgaact catactttcc aggttgtctt tctacgaaaa aattgaaatc 2640
 79 ctgagaaaaa ttgatttttgg taacagtctc aaatcccaag aaaatacagc gctgcaccta 2700
 80 gacaaactga ggcgattgcg taacgcattg gcgcatgcag cacacatgcc acctgatgaa 2760
 81 atcatgaagt tgtgctctga taagtggata gagtcctttg tgctcggata tccaaagtcc 2820
 82 attggcaaag agaaaaatgc acttgaaaat cggctatcac ttctgtgtaa ttactgccac 2880
 83 aggaggcatg tagcaaaaat taagcagctt gcacacgaac tcaaaaatac agagcaagcc 2940
 84 aactaataga gtccagttat acaggtccgt aaatgagccg cctaacaact ggttcaagcc 3000
 85 actcacttcg ctcgctcggg accgcgttcc gcggcccctt aaccaaacgt tgggcaccca 3060
 86 tagaaaaatc ctaatgagaa aactattcat accactaatt ttcgccctgc tatcggagag 3120
 87 cttgatggca tctgaagcgt ataaggacct tgaaacacaa gtaactgaaa aagccagcct 3180
 88 agcagttgcc caaatgaatg acagagcaac tggaaaagtc gactactcgg aagaaagtct 3240
 89 ctatgcagta gaagaaatgg cagcggaagc agctcaatac aaagatcaat tagatccagc 3300
 90 cactgtagac tcgcttactc aagttcttgg aagctatatt cttgaggttg cacatagaaa 3360
 91 gcatggcggc tcttacgttt ggcttgaatc tgaaaactca cctgccttgg tagttggtga 3420
 92 accagagtac aggctagcac tctcaacctt cgccaaggta catggccgac tttctggcga 3480
 93 cgaagcagat aatcttattt tcttctatca aggcttttct gaaaggctta aatcaccatc 3540
 94 tcccggcatg agcgcactct acaaatgaaa cccgagtttg ggggcccaac aatgcgctca 3600
 95 actgccgctc acttcgttcg ctggacgtcc aaaagctacg cttttggccg cccgttagct 3660
 96 taatcgttat gcacaataaa acatgaagac agcactcata tttgtagctc taatctttct 3720
 97 ctctggatgt gacaactatc agtcatgccc tataactgga aaatggaaat ccaacgaaaa 3780
 98 gctaacttta gaaagcatga atgaaaccgg caggataacg gcaaagcaaa gagagatttt 3840
 99 tgagaacggc ttctttggaa aactagaatt agacataaat tgcagtagct tcacaacaat 3900
100 acttgacggc gttaccgaaa cctttaatta cgagatagtc cgccaaacaa aagattccgt 3960
101 caccgttagc tattacagca aagcgctgca aaaacaagtt gaggtcacat ctattatcaa 4020
102 cggaaattgt tactcgacac ctatagagca gttaaatttc aatgagtatt tctgcagagt 4080
103 cgagtagcgc ataacaattg attcaagtcg ttcgcttcgc tcactgcggg accggctaaa 4140
104 gccggcccct taaccaaacg ttaggcaaag gctcaaatgga tcccatattc cataacatcc 4200
105 atagaaacga caaagagatt gagggcgctc atcaacaatg ctcgagcaca atcaatcact 4260
106 tcattgagat ggtcaaaaaa ggggcgagc ccacctatat ggcaaagcta cgttttcttg 4320
107 accctgacaa gtctgaaaaa gaaggtaaga atcatatttt ttatttgtgg ttatctgaag 4380
108 tgctgtacca ccctgcaaca aatttacttt ctggggtatt ttttgaaatc cctgaaggct 4440
```

```
109 ttgaaaagtg gcaccaaata ggccagcgcc taggctttga tccagaagat gtcttttgatt 4500
110 ggatggtaat cgacaaaggt catgctaagg gtgcatacac actaaaggta tcgcgagagc 4560
111 gcttaaccac cgagcaagaa agaaaagatt ttgaccgcta tattggtgtg gcgtcatatg 4620
112 agtagcctaa aattaagcgc tcacgcctca gcctaactac tggttcaagt cactcgcttc 4680
113 gctcgttcgg gaccgcgttc cgcggcccct taaccaaacg ttaggcgcaa gggcaatatt 4740
114 ggtcttcagc accgagtcag gaaacacaat caccgaatca gcgcggtgtt cctgaatcga 4800
115 atggtcgctg acagttgagg ccgttatttg tggccagcaa aggagttgct ttcagagaat 4860
116 gtgcacgtca caaataactt ccggggccaa aaccgaaacg ccgtgcgctc cgccggttaa 4920
117 gctcggcgct ggcatcattt tcggcgctcg gctgggcaat ctaacaattg gctcaagtcg 4980
118 ttcgcttcgc tcactcggga ccggcgaagc cggcccctta gccaaacgtt atgcgagcca 5040
119 ccatgaatag cgaagaatta tacaaaaagg ctatggagtt agagtccaaa tgcgagcata 5100
120 aagcggcaat ttcaacttac aaagaaattg ttaagaaatc taacgatcct cgacacttca 5160
121 tcgcattcgg agtttgcctc caaaaatgtg gtcactgaga gcaatccatc gaggtattag 5220
122 aatcaggaat tgcactgaag cctcactatt gcgagggtga tgctcgtcta tttttagcaa 5280
123 aagcactttt taaatcaggc aaaaaaggcc ttgcgataaa gcaatggcaa catgtatcaa 5340
124 aaatgcaacc tgagtaccca agttatgagt ctgtgcaaaa tgaagccaag aaaatgcttg 5400
125 cacaaaacgc ataacaattg gctcaagccg ctcgctccgc tcactcggac gtccgtaagc 5460
126 tacgcttccg gccgcccctt agccaaacgt taggcaccac atgccctcca tcaagtcagc 5520
127 aagccaatac cagcgcgcaa catcgctcat cttcttagtg tcaggcgctg cttggctatt 5580
128 catcgtgcag tcttcgttgc tgccattgac ggatgtcgcc cgccaagaaa tggtttgcct 5640
129 taatatcgtt cttggtattg cctgtttttgt tataggtagt gcggcaaagc gtcagcgaga 5700
130 atttcgctgc cctgactgtg ggaacgaagt agatcagagc ttacctacag agggtgatgg 5760
131 cgccccactc ctaaggctgt gcaagcactg cgatattcta tggaatgttg gcaagacccc 5820
132 agacagttaa agttaccgcc taacaaatgg ttcaagtcgt tcgcttcgct cactcgggac 5880
133 cggctaaagc cggcccctta accaaacgtt aggcaacagg gggtgacatg acgcaatgtc 5940
134 caaggtgcca gcgcaatctc gcagctgacg agttctatgc tggctctagc aaaatgtgca 6000
135 agggttgcat gacttggcaa aacctaagct acaacgcgaa taaggaaggt catgccaaca 6060
136 ccttcaccaa agcgacattt ttggcgtggt acggcttatc agcacagcgg cattgtgggt 6120
137 attgcggtat atcggaggca ggtttttacat ccttgcacag gactaatcca cgcggctacc 6180
138 acatacagtg tttgggtgtt gatcgctcag attcgttcga aggctattca cctcaaaacg 6240
139 ctcggctcgc ctgtttttata tgcaacagga taaaatcaaa catcttcagc gccagtgaga 6300
140 tggacgttct aggtgaggcc atttcaaaag cgtggcatgg tcgaggaatt gcctaactat 6360
141 tcagtcaagc ggacgcaaac ccgctgcgc ggtctttgcg ccgcttatct caagcgttag 6420
142 atgaataaaa gcctccacac atagccagct ttacgggaac gaagttgatg cgacgcctct 6480
143 ttactgcctt agtattaatt ttgatgattt ctgggtgctc ctccacatca aaaactgaaa 6540
144 gccataaaca gccgccgaac aattcaagcg acacgaccgc catactgaaa tatatattta 6600
145 ctgttaccca tggcatagaa gagggctttg ccaagtcact tgagcaagga agttacacac 6660
146 cagatgaata tatgctcatg cagcaggctt ttagcaatct tgatctaaac agattaacca 6720
147 ctcttttatc acccactta gataaaagca tgaacaccgc agacgtaaaa catttcatga 6780
148 tttttataaa atctactgca ggtaagaact tgctaacagc aggagagtct agcacttctt 6840
149 tttccggcac aatggataga gttcgatcac taccctccga acagcagtcg aaaataaacg 6900
150 agttcttcca tgccagctac acaaaaaaca ctttaacagc catggggggt ccagaggcgg 6960
151 tacgcattgt ttatgcattt ggagtggaat ccatgtgcaa ttacgctatg cgcaataatt 7020
152 tcgaactgta tatttctatt atagaaaaag gcaaatgcca ataaaccaca aaaaacaaag 7080
153 gccggatcga tccatgtgaa catcagcgtt acatctaaca tgtggttcaa gccgctcgct 7140
154 tcgctcactc gggaccggct aaagccggcc ccttaaccaa acgttagagg attacatgcc 7200
155 atcactccaa gaactccaat cgccgattga ctcagcaatc gtgaactcca tgattgagag 7260
156 cactcccgaa acatgaagcc agattattct aacgttggta cgcgaatcta attcctcttg 7320
157 tgtaggtaac tttacacatg agttatccag tcctgagggg catgcaccag ttggtccggc 7380
```

```
158 agagagctta tttgaagaca cttaccaact cgatgagtta ttctacagcc atggtgagcg 7440
159 cttattcacc aaggcaattt atcgggctaa tgcggttggg gacggttggt catatcacgc 7500
160 tgagtttgaa tatgcgtagc atccctctaa caattggttc aagccgctcg cttcgctcgc 7560
161 tcgggatcgg cgaagccggc accttaacca aacgttagag atggtcatga ataaacgcgc 7620
162 actaaccttc gggctactca tagcaattct agctagtatc atttcacaag cgctgcttca 7680
163 tggccaaaag gtaatcgctt ccgacgtggc tttatatact ccatatttcc tatctccaat 7740
164 atttagcatg ctgatacaag ccacatcgat gctggcatgt gccatacctg gactttatgt 7800
165 aggctactta tgcaagaaca agccagcaca acatggagca aaattggggg cagcatatgg 7860
166 aatacttctt ggattaattg tattcgcaat gcgagcttcg acccaattaa cgtagattct 7920
167 aagttaataa tcgcaacatc tgctttaaca caaaaagcaa aatattcagt gcactttgcg 7980
168 ctagttgctc ctgccggcta tcttcttgca aagcatcgtg caaatctcta acaattggtt 8040
169 cagatcgttc gcttcgctca ctgcgggacc ggctgaagcc ggcccttaa ccaaacgtta 8100
170 ggcaactgaa tgatcacctg cattccggca cgtgaattcc tgcgtaaagt atgcggcctg 8160
171 tacgaagcct caactaatgt agttaagttg cgtgtatggg cttgtggata tggcatcgca 8220
172 atggatctaa ctgtcaaagg taaatctgtc cttgtgcgg ttgcgggagt actccgccaa 8280
173 gaggtcgaat gctttgctca aattgggctt ccgaacgtaa ttcagttagt aggcgacaag 8340
174 gcgtcagaga atcaactaaa gctcataggc atggaaccac caatcgaact tcatatctct 8400
175 cgcgaacaga gcaggctcca agttgtaatc ttgtacgagg gtcaggtaaa ggctacatat 8460
176 gtgctgtcag ccgcctaact actggttcaa gtcgttcgct tcgctcactc gggaccggct 8520
177 aaagccggcc ccttaaccaa acgttaggct ttcaatgaaa acagttccag tgaaaatatc 8580
178 agaagtcgaa ctaatagaga gttttgggaa attcctgatc aatcaagact taatcgacta 8640
179 tgaaaattcc cacttcagtg gcgacgacaa ccataatgca gatgtagcct tatctttaaa 8700
180 gccagggaaa tggccaggca ttcaagtcga taaactacac atagaagtaa agtcacacca 8760
181 ctcagaagac tctcaaaaca ccatcaacaa aatattcggc caattactaa aagaaaccgg 8820
182 aaagcgaagc ctcgataaag agaaagagtg cttagctata ttgttccctt acgagcgcgg 8880
183 cgcatggcca ggtcgaaaca acaaaacagt aacaagaatt gaaggtgaag cttattaccg 8940
184 gaggggcttt tcgagaatcg acaaacagac gtttgttaaa tttggtgact tggtcggtgc 9000
185 caaatacatc ctttcctttt ctacagcatc aaacacattg aacgtatttg aatggaaaaa 9060
186 tttcttagat gaggaattca gcccgatgat cagcctaaca aatggttcaa gccgttcgct 9120
187 tcgctcactc gggaccggct aaagccggcc ccttaaccaa acgttagacg caccggaaat 9180
188 tttgcatggg aaaccagaaa atggatttgc agataaacga tacaaaggtt gagtgggttt 9240
189 ctccaatact gaagcaatgg atcagcatca acaaagaata cgtcaagcaa tatgatttca 9300
190 aagactgcct gcactggtat aacgaaaggg caaatataag tgcttttgct ggtgccgttt 9360
191 ggaagtctga aggttttgcg ctggaagaat attcaactaa aaaaggcacc gaagaaaaca 9420
192 gagccaatgg tcgtgtcgac ctatatttct ccaatgacaa cgagcaagcc attgttgaag 9480
193 caaaaatgga atggctctac ttcggaaagc gcacaagact agatttcaaa gaaaaaatag 9540
194 atcgtgtagt tgaaaaagca aagaatgaca taattaacag cctgcatgcc aacccctacg 9600
195 atctaggggct tgggctttcc tttatttgca catactggaa aaagggttat gacgcatccg 9660
196 ccgacatgca agcccttaga gcgcttatgc aaaattataa ctgcgcattt tatgcaatt 9720
197 ttgaaaacgt ccccgacaac gaaattgtta gctcaaaggg caatatctgc aacgctgtga 9780
198 ttttagttgg gacggcgcac agctgaatcg tgtgtgtgcg tctaacaatg cgctcaaagc 9840
199 gctcacttcg ttcgctggga tcggctaaag ccggccccctt agcttaatcg ttagcactag 9900
200 gacttccgac catcatgagt gatagagacg aatttttctgc cccaacaaaa agagcgctag 9960
201 ccgaaaggag tggctttagg tgttcttatc ttggttgctc taatgcaacc ataggggccta 10020
202 gtgaagaatc agaaacagcc gtagcaagaa cggggggtggc gtgtcatata actgccgcag 10080
203 cgcccggcgg aaaaaggtat gacccaacat taagccctac ggaacgaagc tcaatctcga 10140
204 atggtatatg gatgtgccaa acgcattcag ttgaaataga tagagatgag gcccgataca 10200
205 catcgacctt attaaatcac tggaaaaata tatccgagag ccgagcagat tatgcaaaaa 10260
206 atcatggctg ggatattttt gacaaatacc ccttccttca tattgactcg ctagccaaca 10320
```

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/701,626

Input Set : A:\Neb-165a.app
Output Set: N:\CRF3\01302001\I701626.raw

```
207 tagacctggc tcttaccaaa agcccttcct caaatagcct tatcgggaat gccattacag 10380
208 acagctgcct ccctcaacta tggggtaaag agcaatctgt aatcatcaga gacctaataa 10440
209 tagaacttta tcgaaatgcc ttcgatcacg gcgaggctag ctcattcgaa atatccatat 10500
210 cggagcaaaa actagaaata gtttacgatg gcaaaaaatt tgacatcttc caacttcttg 10560
211 accaccagaa tgcaaacggt ggcgccgata ccttgcaaga aattgtagaa aaatatggca 10620
212 gtaactttgt agtcaactat agccacgaag gcaacaataa aataataatt cacaggctct 10680
213 ctgacttta cgcgcttgca ccatccctcc cgtgcgtaat atcactgagt gaatacgatg 10740
214 acaaggccct agagttagac ctggctattt atgagcgctg cggtgcactg tacataattc 10800
215 taccgttgca ttttttgtaga tcagatgtca gggggctaga gtcgcagcta gccgcctttg 10860
216 aacctaatgg aaagccagtt tacattgtag gctcagatgt ggcagagcct acaagaaaag 10920
217 caattataga caggcttccc aacttcacgt tcgtccaaaa gcaatgctaa caatgcgctt 10980
218 aactgtcgct cacttcgttc gctggatagt caaaagctgc gcttttgtct gtccgttagc 11040
219 ttaatcgtta ggcgcaagga gggaccgtga ctgaaactga gaaaatggtg ggtaagttcg 11100
220 tcagcggttt tggcgggcag agataccgag aaattttga agtcctcgaa tccagtaacc 11160
221 ttcgcccact gggcaagtca aatactgaaa cattgctatt tcagcttcga ggggctgata 11220
222 gtgaaatgct agatattttt gcctttcgct tggggccgcc gccagtaatt tcgtttccca 11280
223 aatcatattg gctaggtcgc cccagtgaat taagcgctca tctatccaat ttttcattct 11340
224 cggaaaagcc agccataaca ggcccggttt ctgactcaca gtattcggca ggccaggtgg 11400
225 aaatcacccg ctctactcat gagaggatta ttgaggtttg caaccgtgtc tgtgcttccc 11460
226 tgcaataagc gcctaacaac tggttcaaat cgctcgctcc gctcgctggg accggcatag 11520
227 ccggcccctt aaccaagcgt tagatgcaaa taacttgagg ggcacatgca agactttggg 11580
228 tccagacgaa atgcatcatt agaggacagg gctgcggctg agtctgttat tgaacgtgtt 11640
229 tatcttgcga tacagcagct ttgcacagag actggtgacg taagaaatcg gcttcaaata 11700
230 gccgttatga ctctattgcc ccttcaggcg cgtaacttcc ccattgcgtt gcagcaagac 11760
231 ttcgattgga ttgtcagaga atcaaccaaa tacaaatcac catatccgca gtttcggggc 11820
232 gaccttgaag caacgatgat gcgaataagg aactcaactg ggcaaaaaat cgcgcaaaga 11880
233 attttcaata tttactcgtc gctacaagac attcgaggtt ttcccctgct tgaatacagg 11940
234 gcaatagatg agtaagcatc taacaattgg ttaaaaccgt tcgcttcgct cactgggacc 12000
235 ggctaaagcc ggcccctaa ccaaacgtta ggtaaccaag ggaaattcac ttgagttgtt 12060
236 atgtattggg cacaaacaac cgccattaaa ggacggtttt atagtaaatt tcatcggact 12120
237 gttgaactaa aatgcttata cgctttgctc tactacttgc tgttatgctc ctcgctgcat 12180
238 gctcgtcaaa gcaaaatcca acgccgaagt gtactgccag cgtcccccg ccctctttac 12240
239 ccgaaacatc cacagtatgc ctagggaaa gatgtaattg ggaggtgcta tttccgtcag 12300
240 gaaaataccc tgcatccaca gaaggctgca gggcgcctgt ggtgcagaac cagccttctt 12360
241 cctacccgcg agaagcactt gatcagtgta ttgaggggta cgcttgggta gcggtttttc 12420
242 tgaatgccga cggggtccaa acatcagcaa aggtacttca atcatcgaat aaaattttcg 12480
243 acagaaatgc cttgctcagg gccagtaata tatttttga gcctatgaaa tgtcaatccg 12540
244 agcgttatga ttccgttgtt ctgatgccat taaactaccg catactcccc tagtagcggg 12600
245 attgatcctt acaaaattca ctacttacgt ccaagttgaa gtaggcagtt taacaactgg 12660
246 ttcaagccgc tcgcttcgct cactcgggac cggctaaatt cggcccctta ggcaaacgtt 12720
247 aactatcaga agggcggttg atgtcaagat ttgcgctcgc gttgattcac ggagtaccaa 12780
248 cgggttttct tgtcatttgt acttttgtttg tctgtttcat ctacctcaac cgattcgaga 12840
249 aagttggagg atactcagac gggtgggggtt ttgttggaag agttgtctgc gcatctatag 12900
250 ctatggtttt cgtgtccgca gttggccatc ttcttattga agcggcagtc aactgggggc 12960
251 tgcagcagct tggttatgag ctgccaaact atgaaaaaag aaggacttgt agtagctgca 13020
252 agccgagcac tccaggtgac tacatgttcg gcttgctcct cggggttgtg cttggcgccg 13080
253 gctcggcaat ttggctctgg acgcgcctgg cgctccgata tgcgctgttt cgcggcgaaa 13140
254 actgatagct gaaccttcca tcgaggagat gcaaaagcgc tgctgcgcgc catctacaaa 13200
255 gacccgaagc acctcatcca ggcgctctca gcccgagcct gactggctgt ggctatcaac 13260
```

```
<210> 4
<211> 42143
<212> DNA
<213> Unknown

<220>
<223> Description of Unknown Organism: Genomic DNA of
      Pseudomonas Alcaligenes NEB #585 (ATCC 55044)

<400> 4
atcgatcagc cagactttc gcacacgggc ggaccttggg cgagtcagcg ctatggttgg 60
ccgctgtggg ttgtcagtgc ccgtacgcgc aatctgtttc tttcgcaggg catgtccggc 120
tgggcgttcc ggcccgttct ggtcaccgac tcggctctct atgagcgcta tctcgctcta 180
agtcaggaac tttgcgcact gcttcgtgat gcaccgcaga gcaagctcga agaccgtgat 240
tggtaagcgg gggctattcg atcagtctcg gagcgaccaa actccagaaa cgacaaggcc 300
ctgaaaaaaa agcagggctt cgtctttgcg ggcgaatgga atcggacctc tttccgcctc 360
tgcatgtaac tggtctttgt ttgccaaatc tgcctatctc atgccggcca tgttggccag 420
tgcctgcatc atttggcctt tggtttcgac acttttttcga cagccctgct agacatccct 480
ccctctgccc tcgtaacttc tgttccgatg gtgtcgcttg cactatggt cttgtcgagt 540
gtcgctttc atccagccta atgccgcgat tgcctcgctg agctgtagct gaatcaagga 600
cttagcggac gacaaggaat gttatgcgaa acatgtggcg gaataaatta cgccgcatgt 660
ttcgtctact tatagttagg ctacatatga gaatcagcgc agaccagctt gctcaagaat 720
cactgactga gttcggcgtg ctggcggcta agcttctggc aacgcgagag cttagccagt 780
tgtccgagaa gtttgggtat gcactggcct tcggaaggga accggcggct gccatagctg 840
aggaccttgc taggtgcttg tgcggacaaa atgcttcgcc ggcatctgaa taccccaaaa 900
tcaccgttaa gtatttcaag gaaaacgaaa gtagtctgtt ggcactcgta gagtgttatg 960
tacaaatgac cgcaagcgca aacattcttt tagagctggt tgccgcacga aatggagagg 1020
caataaatct gtatctagaa ggcttgagtg ttgtagccta acaatgcgct caaagcgctc 1080
acttcgttcg ctgggaccgg cgaagccggc cccttagctt aatcgttaga aaccatcatg 1140
gataactggt acaacaccat cgaataccaa acccatgtag ccgaaaaact agaggcactt 1200
ggagaaacaa agtacgaccg cgaggcttat gaattcgcgc tagaggcata ccagtatgcg 1260
cctgaatatc atgaaaatat tcccacgccg cctctcaatc ttgggctcgc gtaccatgta 1320
agcgccttca actttgcaca ctgctatgta cttcacgcta aagaagtgtt tgaagctcca 1380
aaagacacac tgagctcctg gggcgtattt tcctcaacgg acattggtga aattgtttat 1440
ggtttagtcc gtattggctt gctggaccaa ggccccgaag acaaaaaaga gcagtttgaa 1500
gggttgtttt taatcaccga cgtgctgtga tgtcttctaa ctactggttc aagtcgttcg 1560
cttcgctcac tcgggaccgg ctaaagccgg ccccttaacc aaacgttagc cacctcacga 1620
agatttggag cccgcgtgaa caaagtcgat acaaacaaaa ttaaaacgga tttttcggca 1680
cgaattgatg aaaaaagagc gtggtttgat cgtatggcta cgcttataag cgggacaaac 1740
accgagttaa ccgaccttaa ttttctttgc gagaactata taacatcaat atacgtagag 1800
ctcgaatgct taatatcaga tttatttcat ggctacataa ataacaacaa caagacctac 1860
atggcgcaca ttcaatcaaa aatcaagaac tccataactg acaagtactc tgcatggcac 1920
gccacccata caacattcgc aggtccagag catattaatt cagcacagct cagcacgctc 1980
cttgatccaa caagctggaa catcacattt aaagacgttt ccgcaatgaa agtacgagca 2040
aaggaatacc tttcctcagt acacgaaaaa agattttcag gtatatctgc atccgatgga 2100
gctcttattg atgccgcaca tgcaatcaga aattgcattg cacacaacag cgaaagctcc 2160
agaaaggtta tgaacaccaa aattaaaagc ttaattacag gcccagcttg ctcaaatgtc 2220
ggccttgaac tcaccacaaa tagtgtgacc aaaataggaa agtatctccg tgcaaatgct 2280
cagcaaagca tgcgagtgct gatttactca gatcgaataa aatctatcgg cctaagctta 2340
taagtgtggg ctaacaatgc gctcaactgt cgctcacttc gttcgctgga cagccaaaag 2400
ctacgctttt gtctgcccgt tagcttaatc gttaggaggc tctgcatgac tcgtgcaaca 2460
gacaggttcg aagagcttct gcaatcacat gagttctcag gcatattat tcgttgggtt 2520
gcgatattcg aaggccgtct tgacggtgtg ttatcagttc attttctgg acttgaaagc 2580
acctatgaat tctacgaact catactttcc aggttgtctt tctacgaaaa aattgaaatc 2640
ctgagaaaaa ttgattttgg taacagtctc aaatcccaag aaaatacagc gctgcaccta 2700
gacaaactga ggcgattgcg taacgcattg gcgcatgcag cacacatgcc acctgatgaa 2760
atcatgaagt tgtgctctga taagtggata gagtcctttg tgctcggata tccaaagtcc 2820
```

```
gctcggcaat ttggctctgg acgcgcctgg cgctgcgata tgcgctgttt cgcggcgaaa 13140
actgatagct gaaccttcca tcgaggagat gcaaaagcgc tgctgcgcgc catctacaaa 13200
gacccgaagc acctcatcca ggcgctctca gcccgagcct gactggctgt ggctatcaac 13260
acctcttcga taccactacc cgccagaaac gacaaagccc tgcaaaaagc agggctttgt 13320
ctttggggat ctggagcggg cgaagggaat cgaaccctcg tcatgagctt gggaagctca 13380
ggtaatgcca ttatacgacg cccgctcggg cggctgactt tttaccagaa tcgcccggga 13440
aggtgaagcc gggcgcgcgt cttgcgcccg ttttattgcc gggcgcttca tagcgccacg 13500
gcccgtggct ctcgttccac gctgcgtgcg tggccctgcg tgggtgccag caggaaggcc 13560
agcagggcat cgcgggtctg catccaggcg gccttgtgtt ccatgtcgag gaagtggccg 13620
gcctgggcga tggtgcggaa ctcgcagtgg cgcacgtact gggtgaacag gcgcgcgtca 13680
gccggggtgg tgtactcgtc ccactcgccg ttgacgaaca gcagcggtat ctcgatctgc 13740
ccggcgaagc tgacgcagga gcgcccgccg ttgttcagca cggtttccac gtggtgactc 13800
atttgctcat attcatagcg ctccaggccg gtgacgtgtc gatggttgta gcgcttgaac 13860
agcgagggca ggtgcttgcc gatggtgccg ttgagcacca tgccgatgct ctcgcggtcg 13920
cactcgcgca tcaccaccag gccggcgcgc aggtagccga gcatggcgct gttgacgatc 13980
ggcgagaagg agttgatcac cgcacgctcg atccgcgatg gacgccgggc cagcgcctgg 14040
agggtggcga tgccgcccca ggagaaagga cagcacgctg ttcgcaggcg aaaatgttcg 14100
accagctcca ggaagatgtc ggcttctttc ctcgcsgctg aaggrbsdna agcttctggt 14160
acgaacctgg gggcgctccg gcacgcacaa gggcatcgac atcttcgccc gccagggcac 14220
cccggtgctc gcccccagct acggcatcgt ggtgtttcgc gacgagctcg acatgggcgg 14280
caaggtactg ctgatgctcg gccccaaatg gcgcctgcac tacttcgccc acctcgacag 14340
ctacagcgcc ctgcccggcc aacccgtact tcccggcgcc ccactcggca cggtaggcag 14400
caccggcaac gcccagggca agccgcccca tctgcactac tcgatcgtca ccctgttgcc 14460
ctatccctgg cgctgggaca acagcactca gggctggaag aaaatgttct acctcgaccc 14520
cacgccaatg ctgaacgaag cggcagtaga cagccgaaaa accagccagt agcgtcgcag 14580
gggaatgcac caccggtctt gcccgatccg cctgtccttt taccaatcgc agaagagtcg 14640
cttttgtcga atcgcctgtg aggaaaaaca aggacttgct ggacgacaag gaacgttatg 14700
cgacacaagt ggcggaataa attacgccat ttgtgtcgtc tacttatagt tatatgctga 14760
tctagatatg aagtacaaaa acataaaatc agcaatccac aatttcgggc acagctttgt 14820
aagctcagtg aactatgttg accatgattt cgttgccgac gaaattggga agattcacaa 14880
gaaaggctat gatattgaaa taaactggct tacaagggag ttcaagcccg ctcagcttga 14940
gtcagagaga ataaaaaaat caattggtta ttggggtgac aacctaaaga aacattgtgc 15000
```

due to size of Sequence 4, only the first error in that sequence shown. Identical error appear elsewhere in the sequence.

→ see item 10 on Error Summary Sheet

L:12 M:270 C: Current Application Number differs, Replaced Current Application Number
L:13 M:271 C: Current Filing Date differs, Replaced Current Filing Date
L:897 M:258 W: Mandatory Feature missing, <221> not found for SEQ ID#:4
L:897 M:258 W: Mandatory Feature missing, <222> not found for SEQ ID#:4
L:897 M:340 W: (46) "n" or "Xaa" used: Feature required, for SEQ ID#:4
L:1019 M:258 W: Mandatory Feature missing, <221> not found for SEQ ID#:4
L:1019 M:258 W: Mandatory Feature missing, <222> not found for SEQ ID#:4
M:340 Repeated in SeqNo=4
L:1141 M:258 W: Mandatory Feature missing, <221> not found for SEQ ID#:4
L:1141 M:258 W: Mandatory Feature missing, <222> not found for SEQ ID#:4